



Codebook

V 9 - April 2019

Suggested citation: Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Valeriya Mechkova, Johannes von Römer, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 2019. "V-Dem Codebook v9" Varieties of Democracy (V-Dem) Project.

Main coders of factual (A/A*) data: Frida Andersson, Vlad Ciobanu, Lee Cojocaru, Lisa Gastaldi, Valeriya Mechkova, Selemon Negash, Moa Olin, and Laura Saxer.

Other coders of factual (A/A*) data: Tove Ahlbom, Pelle Ahlin, Eric Bader, Julia Bianco, Adam Bilinski, Solveig Bjørkholt, Sophie Cassell, Felix Dwinger, Kristian Frederiksen, Ionut Gusoi, Dominik Hirndorf, Sune Hellegaard, Bernardo Isola, Talib Jabbar, Haakon Jernsletten, Eva Karlsdottir, Berker Kavasoglu, Thomas Krauchi, Joshua Krusell, Ole Martin Laegreid, Frederik Lasserre, Vilde Lunna Djuve, Claudia Maier, Swaantje Marten, Mark Patteson, Livia Radaeski, Hayley Rasuchine, Janina Schlect, Tove Selnes, Katharina Sibbers, Konstantinos Skenteris, Andrew Slivka, Tolga Tan, Marcus Tannenberg, and Philipp Tönjes.

Funders: We are very grateful for our funders' support over the years, which has made this venture possible. To learn more about our funders, please visit: <https://www.v-dem.net/en/v-dem-institute/funders/>

For data enquires: contact@v-dem.net

- **Appendix D: Post-Survey Questionnaire**

This survey is completed by all coders. Data from the survey is not included in the V-Dem Dataset but may be provided on request (subject to review and ethics approval).

- **Appendix E: Comments section**

This section lists how the request for comments were phrased in each survey. Comments made by coders are not included in the V-Dem Dataset but may be provided on request (subject to review and ethics approval).

- **Appendix F: Changes Between Previous Versions of the Dataset**

1 Explanatory Notes

1.1 New in Version 9 compared to Version 8

- **New country**

Malta (Coded: 1900–2018).

- **New thematic areas**

Exclusion, Legitimation and Digital Society.

- **The Digital Society Survey**

The Digital Society Survey, designed by the Digital Society Project, contains questions pertaining to the political environment of the internet and social media. The data collected through expert-coded surveys provides information on topics related to coordinated information operations, digital media freedom, online media polarization, social cleavages as well as state internet regulation capacity and approach.

For more information, please visit <https://www.digitalsocietyproject.org>

- **New indices**

- Exclusion by Gender (v2xpe_exlgender)
- Exclusion by Social Group (v2xpe_exlsocgr)
- Exclusion by Urban-Rural Location (v2xpe_exlgeo)
- Exclusion by Socio-Economic Group (v2xpe_exlecon)
- Exclusion by Political Group (v2xpe_exlpol)
- Hereditary dimension index (v2x_ex_hereditary)
- Military dimension index (v2x_ex_military)
- Ruling party dimension index (v2x_ex_party)
- Direct election dimension index (v2x_ex_direlect)
- Confidence dimension index (v2x_ex_confidence)

- **New indicators**

- Regime information (v2reginfo)
- Regime end type (v2regendtype)
- Ideology (v2exl_legitideol)
- Person of the Leader (v2exl_legitlead)
- Performance legitimation (v2exl_legitperf)
- Rational-legal legitimation (v2exl_legitratio)
- Exclusion variables:
 - * Power distributed by urban-rural location (v2pepwrgeo)
 - * Gender equality in respect for civil liberties (v2clgencl)
 - * Political group equality in respect for civil liberties (v2clpolcl)
 - * Urban-rural location equality in respect for civil liberties (v2clgeocl)
 - * Access to public services distributed by social group (v2peapssoc)
 - * Access to public services distributed by gender (v2peapsgen)
 - * Access to public services distributed by socio-economic position (v2peapsecon)
 - * Access to public services distributed by political group (v2peapspol)
 - * Access to public services distributed by urban-rural location (v2peapsgeo)
 - * Access to state jobs by social group (v2peasjsoc)
 - * Access to state jobs by gender (v2peasjgen)

- * Access to state jobs by socio-economic position (v2peasjsoecon)
- * Access to state jobs by urban-rural location (v2peasjgeo)
- * Access to state jobs by political group (v2peasjpol)
- * Access to state business opportunities by social group (v2peasbsoc)
- * Access to state business opportunities by gender (v2peasbgen)
- * Access to state business opportunities by socio-economic position (v2peasbecon)
- * Access to state business opportunities by political group (v2peasbepol)
- * Access to state business opportunities by urban-rural location (v2peasbegeo)
- Digital society variables:
 - * Government dissemination of false information domestic (v2smgovdom)
 - * Government dissemination of false information abroad (v2smgovab)
 - * Party dissemination of false information domestic (v2smpardom)
 - * Party dissemination of false information abroad (v2smparab)
 - * Foreign governments dissemination of false information (v2smfordom)
 - * Foreign governments ads (v2smforads)
 - * Government Internet filtering capacity (v2smgovfilcap)
 - * Government Internet filtering in practice (v2smgovfilprc)
 - * Government Internet shut down capacity (v2smgovshutcap)
 - * Government Internet shut down in practice (v2smgovshut)
 - * Government social media shut down in practice (v2smgovsm)
 - * Government social media alternatives (v2smgovsmalt)
 - * Government social media monitoring (v2smgovsmmon)
 - * Government social media censorship in practice (v2smgovsmcenprc)
 - * Government cyber security capacity (v2smgovcapsec)
 - * Political parties cyber security capacity (v2smpolcap)
 - * Internet legal regulation content (v2smregcon)
 - * Privacy protection by law exists (v2smprivex)
 - * Privacy protection by law content (v2smprivcon)
 - * Government capacity to regulate online content (v2smregcap)
 - * Government online content regulation approach (v2smregapp)
 - * Defamation protection (v2smlawpr)
 - * Abuse of defamation and copyright law by elites (v2smdefabu)
 - * Online media existence (v2smonex)
 - * Online media perspectives (v2smonper)
 - * Online media fractionalization (v2smmefra)
 - * Online harassment groups (v2smhargr)
 - * Other online harassment groups (v2smhargrtxt)
 - * Use of social media to organize offline violence (v2smorgviol)
 - * Average people's use of social media to organize offline action (v2smorgavgact)
 - * Elites' use of social media to organize offline action (v2smorgelitact)
 - * Types of organization through social media (v2smorgtypes)
 - * Party/candidate use of social media in campaigns (v2smcamp)
 - * Arrests for political content (v2smarrest)
 - * Polarization of society (v2smpolsoc)
 - * Political parties hate speech (v2smpolhate)

- **Modified indices and indicators**

- The "Elected officials index" (v2x_elecoff) has been modified.
- All multiple selection questions coded by country experts are now dichotomized and aggregated by mean.

1.3 Variable Types

The V-Dem Codebook divides variables into the following variable types:

- **Type A*:** **Variables coded by Research Assistants**
This data is based on extant sources and is factual in nature. Country Experts indicate their confidence for this pre-coded data.
- **Type A:** **Variables coded by Project Managers and Research Assistants**
This data is based on extant sources and is factual in nature.
- **Type B:** **Variables coded by Country Coordinators or Research Assistants**
The coder is typically a graduate student or recent graduate from the country in question. These variables are factual in nature.
- **Type C:** **Variables coded by Country Experts**
A Country Expert is typically a scholar or professional with deep knowledge of a country and of a particular political institution. Furthermore, the expert is usually a citizen or resident of the country. Multiple experts (usually 5 or more) code each variable. More information about the Country Experts can be found in the *V-Dem Methodology* document.
- **Type A,C:** **Variables coded by Country Experts and crosschecked by Research Assistants**
- **Type D:** **Indices**
Variables composed of type A, B, or C variables. This data may be accomplished by adding a denominator (*e.g.*, per capita), by creating a cumulative scale (total number of...), or by aggregating larger concepts (*e.g.*, components or indices of democracy).
- **Type E:** **Non-V-Dem variables**
If we import a variable from another source without doing any original coding, except for perhaps imputing missing data, it is not considered a V-Dem product. These variables are found in the sections of the Codebook labeled *Background Factors* and *Other Democracy Indices and Indicators*. If, however, we gather data from a number of sources and combine them in a more than purely mechanical fashion (requiring some judgment on our part), we regard this as a V-Dem product and classify it as type A, B, or C. All "E" variables, except those drawn from sources that have more than 30 variables, are included in the codebook. For this reason, the following examples are not included; Archigos (Goemans et al.), BDM (Bueno de Mesquita et al.), Henisz/POLCON (2000, 2002), Miller (Democratic Pieces), Performance of Nations (Kugler and Tammen), PEI (Norris et al.), PIPE (Przeworski et al.) and QoG (Quality of Government). For these and for similar cases, we ask users to consult separate codebooks, as listed above.

1.4 Structure of Aggregation

The V-Dem conceptual scheme recognizes several levels of aggregation. The table below shows the structure of aggregation for one of our high-level indices: The V-Dem Electoral Democracy Index. This index consists of five sub-components (each of these sub-components being indices themselves built from a number of indicators) that together capture Dahl's seven institutions of polyarchy: freedom of association, suffrage, clean elections, elected executive, and freedom of expression and alternative sources of information.

Note: *Appendix A* includes a table with a complete hierarchy of our democracy indices, democracy component indices, democracy sub-component indices, and indicators, as well as the hierarchy of related concept indices.

1.5 Variable Versions and Suffixes

The V-Dem Dataset contains several versions of the variables coded by country experts (type C variables).

- **Model Estimates**

"Model Estimates" — Measurement Model Output:

This version has no special suffix (e.g. v2elmulpar). This version of the variables provides country-year (country-date in the alternative dataset) point estimates from the V-Dem measurement model (see Pemstein et al. 2019). The measurement model aggregates the ratings provided by multiple country experts and, taking disagreement and measurement error into account, produces a probability distribution over country-year scores on a standardized interval scale (see the *V-Dem Methodology* document). The point estimates are the median values of these distributions for each country-year. The scale of a measurement model variable is similar to a normal ("Z") score (e.g. typically between -5 and 5, with 0 approximately representing the mean for all country-years in the sample) though it does not necessarily follow a normal distribution. For most purposes, these are the preferred versions of the variables for time series regression and other estimation strategies.

"Model Estimates Measure of Uncertainty" — Measurement Model Highest Posterior Density (HPD) Intervals:

This version has the suffixes: "codelow" and "codehigh" (e.g. v2elmulpar_codelow and v2elmulpar_codehigh). These two kinds of variables ["code low" and "code high"] demarcate the interval in which the measurement model places 68 percent of the probability mass for each country-year score, which is approximately equivalent to one standard deviation upper and lower bounds. If the underlying posterior distribution is skewed, the HPDs reflect this with unequal distances between the point estimate and the high and low estimates. We also provide a standard calculation for standard deviation which is marked with the suffix "sd" (e.g., v2elmulpar_sd). The SD might be used to compute the standard frequentist confidence intervals.

- **Original Scale (*_osp)**

"Original Scale" — Linearized Original Scale Posterior Prediction:

This version has the suffix "_osp," (e.g. v2elmulpar_osp). In this version of the variables, we have linearly translated the measurement model point estimates back to the original ordinal scale of each variable (e.g. 0–4 for v2elmulpar_osp) as an interval measure. The decimals in the _osp version roughly indicate the distance between the point estimate from the linearized measurement model posterior prediction and the threshold for reaching the next level on the original ordinal scale. Thus, a _osp value of 1.25 indicates that the median measurement model posterior predicted value was closer to the ordinal value of 1 than 2 on the original scale. Technically, it calculates the sum of the posterior probabilities that the estimate is in a particular category: If a particular country-year-variable has a probability of 90% to be in category "4", a 10% probability of being in category "3", and 0% probability of being in categories "2", "1", and "0", the result is a value of 3.9 ($4*0.9 + 3*0.1 = 3.6+0.3$). Since there is no conventional theoretical justification for linearly mapping ordinal posterior predictions onto an interval scale, these scores should primarily be used for heuristic purposes. Using the "Ordinal Scale" estimates—or incorporating the properties of ordinal probit models into the estimation procedure—is thus preferable to using the _osp estimates in statistical analyses. However, since the _osp version maps onto the coding criteria found in the V-Dem Codebook, and is strongly correlated with the Measurement Model output (typically at .98 or higher), some users may find the _osp version useful in estimating quantities such as marginal effects with a clear substantive interpretation. If a user uses _osp data in statistical analyses it is imperative that she confirm that the results are compatible with estimations using Measurement Model output.

"Original Scale Measure of Uncertainty" — Linearized Original Scale HPD Intervals:

This version has the suffixes – "codelow" and "codehigh" (e.g. v2elmulpar_osp_codelow and v2elmulpar_osp_codehigh). We estimate these quantities in a similar manner as the Measurement Model Highest Posterior Density Intervals. These two variables ["code low" and "code high"] demarcate the interval in which the measurement model places 70 percent of the probability mass for each country-year score, which is approximately equivalent to one standard

deviation upper and lower bounds. If the underlying posterior distribution is skewed, the HPDs reflect this with unequal distances between the point estimate and the high and low estimates. We also provide a standard calculation for standard deviation which is marked with the suffix "sd" (e.g., v2elmulpar_sd). The SD might be used to compute the standard frequentist confidence intervals.

- **Ordinal Scale (*_ord)**

"Ordinal Scale" — Measurement Model Estimates of Original Scale Value:

This version has the suffix "_ord" (e.g. v2elmulpar_ord). This method translates the measurement model estimates back to the original ordinal scale of a variable (as represented in the Codebook) after taking coder disagreement and measurement error into account. More precisely, it represents the most likely ordinal value on the original codebook scale into which a country-year would fall, given the average coder's usage of that scale. More specifically, we assign each country-year a value that corresponds to its integerized median ordinal highest posterior probability category over Measurement Model output.

"Ordinal Scale Measure of Uncertainty" — Original Scale Value HPD Intervals:

This version has the suffixes - "codelow" and "codehigh" (e.g. v2elmulpar_ord_codelow and v2elmulpar_ord_codehigh). We estimate these values in a similar manner as the Measurement Model Highest Posterior Density Intervals. These two variables ["code low" and "code high"] demarcate the interval in which the measurement model places 70 percent of the probability mass for each country-year score, which is approximately equivalent to one standard deviation upper and lower bounds. If the underlying posterior distribution is skewed, the HPDs reflect this with unequal distances between the point estimate and the high and low estimates. We also provide a standard calculation for standard deviation which is marked with the suffix "sd" (e.g. v2elmulpar_sd). The SD might be used to compute the standard frequentist confidence intervals.

- **Number of Coders per Country, Variable and Year/Date (*_nr)**

The number of V-Dem Country Experts (regular coders, bridge- and lateral coders) who provided data on country, variable and year. V-Dem's methodology is based on the assumption that we have a minimum of five Country Experts for every single country-variable-year. Sometimes, however, we end up with fewer than five Country Experts. From v7 of the Country-Year, and the Country-Date type datasets, we provide all data we have for full transparency. By providing the number of Country Experts for each country-variable-year/date, we suggest that users primarily base analyses on observations based on five or more coders. We strongly advise against using observations based on three or fewer coders. This concerns all C type variables.

1.6 Aggregation

C-variables, ratio/percentage variables, and High-Level/Mid-Level Democracy indices are aggregated from the country-date level to the country-year level by the day-weighted mean. Ordinal A-variables and A*-variables are aggregated by taking the last observation in the year with one exception: event-specific dichotomies or event-specific ordinal variables, which mostly concerns elections or election related data, are aggregated by max (meaning the highest observed value for a given year is retained) to reflect that an "event" of the coded type occurred within the year.

1.7 Variable Tags

Every variable has a name and a tag. The tag consists of three or four parts and has the following structure:

Prefix + Index (if V-Dem index) + Section + Abbreviated title

While the prefix specifies the variable type, the letters that follow indicate which section the variable belongs to.

votes studied (popular initiatives $\times 1.5$, referendums $\times 1.5$, plebiscites, and obligatory referendums). Each type of popular vote receives a maximum score of two resulting from the addition of two terms (easiness of initiation and easiness of approval), where each term obtains a maximum value of one. As we are studying four types of popular votes, the minimum value is 0, and the maximum is 8. In the *v2xdd_dd* all scores are normalized to range between 0 and 1. For an elaboration of the weighting factor of each component, see: Altman, David. 2017. The index is aggregated using this formula:

$$v2xdd_dd = (v2xdd_i_pi) \times 1.5 + (v2xdd_i_rf) \times 1.5 \\ + (v2xdd_i_pl) + (v2xdd_i_or)$$

Regarding each type of citizen initiated popular vote (i.e., popular initiative), the ease of initiation is measured by (a) the existence of a direct democracy process (*v2ddlexci*), (b) the number of signatures needed (*v2ddsigpci*), and (c) time-limits to circulate the signatures (*v2ddsigdci*). Easiness of approval is measured by the surface of the polygon determined by (a) participation quorum (*v2ddsigdci*), (b) approval quorum (*v2ddpartci*), and (c) supermajority (*v2ddspmc*). The resulting score is then multiplied with (d) district majority (*v2ddadmci*). Consequences are measured by (a) the legal status of the decision made by citizens (binding or merely consultative) (*v2ddlexci*), and (b) the frequency and degree of success with which direct popular votes have been held in the past (*v2ddthreci*). The index is aggregated using this formula:

$$v2xdd_dd = [(IF v2ddlexci > 0, 1, 0) \times (1 - (v2ddsigpci)) \times (IF v2ddsigdci \\ = 0, 1, .5 + (2 \times v2ddsigdci/365)) + \\ (v2ddsigdci) \cap (v2ddpartci) \cap (v2ddspmc)] \\ \times (0.5 + ((100 - v2ddadmci)/100))/2] \\ \times (IF v2ddlexci = 2, 1, IF v2ddlexci = 1, 0.75, v2ddlexci = 0, 0) \\ \times (IF years\ since\ last\ successful\ event < 6, \\ v2ddthreci = 1, afterwards\ decreases\ by\ 0.06\ units\ per\ year \\ until\ 0.1; if\ the\ event\ was\ not\ successful \\ during\ the\ first\ years\ v2ddapprci \\ = 0.9, afterwards\ decreases\ by\ 0.1\ units\ per\ year\ until\ 0.1)$$

In case the vote originates from above (i.e., authorities), there is no need to account for *v2ddsigpci* and *v2ddsigdci*. For an elaboration of the interaction among quorums, (*v2ddsigdci*) \cap (*v2ddpartci*) \cap (*v2ddspmc*), see Altman, David. 2017.

Citation: *V-Dem Codebook* (see suggested citation at the top of this document).

Data release: 5, 7-9. Release 7 modified: v2pepwrses, v2pepwrsoc and v2pepwrigen now form a separate subcomponent index.

Aggregation: The index is formed by taking the point estimates from a Bayesian factor analysis model of the indicators for particularistic or public goods v2dlencmps, means tested vs. universalistic welfare policies v2dlunivl, educational equality v2peedueq and health equality v2pehealth.

Citation: Sigman *et al.* (2015, *V-Dem Working Paper Series* 2015:22); *V-Dem Codebook* (see suggested citation at the top of this document).

Years: 1900-2018

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology*).

Citation: Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 1789-2018

3.1.127 Subnational elections held (C) (v2elffelrbin)

Project Manager(s): Kelly McMann

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Are subnational elections held?

Responses:

0: No.

1: Yes.

Scale: Dichotomous, converted to interval by the measurement model.

Notes: As of December 2014, the former category "5" from variable v2elffelr is recoded as a separate variable (v2elffelrbin). If a coder chose the 5th category in the original question, it receives 0 in the new "v2elffelrbin" variable (corresponding to the answer, no, there were no subnational elections); otherwise it receives 1 (yes, there are subnational elections held). The resulting series of 0-1 country-coder time-series is run in the measurement model, which calculates the final value of v2elffelrbin while taking into account background coder characteristics.

Data release: 3-9.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology*).

Citation: Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 1789-2018

3.1.128 Subnational election unevenness (C) (v2elsnlff)

Project Manager(s): Kelly McMann

Additional versions: *_osp, *_ord, *_codelow, *_codehigh, *_sd, *_mean, *_nr

Question: Does the freeness and fairness of subnational elections vary across different areas of the country?

Clarification: Subnational elections refer to elections to regional or local offices, as specified above.

Responses:

0: Yes. Subnational elections in some areas of the country are significantly more free and fair (or, alternatively, significantly less free and fair) than subnational elections in other areas of the country.

1: Somewhat. Subnational elections in some areas of the country are somewhat more free and fair (or, alternatively, somewhat less free and fair) than subnational elections in other areas of the country.

2: No. Subnational elections in most or all areas of the country are equally free and fair (or, alternatively, equally not free and not fair).

Ordering: If answer is "2", skip remaining questions in this section.

Scale: Ordinal, converted to interval by the measurement model.

Data release: 1-9.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology*).

Citation: Pemstein *et al.* (2019, *V-Dem Working Paper Series* 2019:21)

Years: 1789-2018

3.1.129 Subnational election area less free and fair name (C) (v2elsnless)

Project Manager(s): Kelly McMann

Question: In which areas of the country are subnational elections significantly *less* free and fair than the country average for subnational elections?

Clarification: If providing names of all the relevant territorial units is not possible, use broad categories (for example, "the North").

Responses:

Text.

Data release: 3-9. Available upon request, subject to review and approval.

Citation: *V-Dem Codebook* (see suggested citation at the top of this document).

Years: 1900-2018

3.1.30 Subnational election area less free and fair characteristics (C) (v2elsnlfc)

Project Manager(s): Kelly McMann

Additional versions: * _nr

Question: How would you describe the areas of the country in which elections are significantly *less* free and fair?

Clarification: Choose all that apply.

Responses:

- 0: Rural. (0=No, 1=Yes) [v2elsnlfc_0]
- 1: Urban. (0=No, 1=Yes) [v2elsnlfc_1]
- 2: Areas that are less economically developed. (0=No, 1=Yes) [v2elsnlfc_2]
- 3: Areas that are more economically developed. (0=No, 1=Yes) [v2elsnlfc_3]
- 4: Inside the capital city. (0=No, 1=Yes) [v2elsnlfc_4]
- 5: Outside the capital city. (0=No, 1=Yes) [v2elsnlfc_5]
- 6: North. (0=No, 1=Yes) [v2elsnlfc_6]
- 7: South. (0=No, 1=Yes) [v2elsnlfc_7]
- 8: West. (0=No, 1=Yes) [v2elsnlfc_8]
- 9: East. (0=No, 1=Yes) [v2elsnlfc_9]
- 10: Areas of civil unrest (including areas where insurgent groups are active). (0=No, 1=Yes) [v2elsnlfc_10]
- 11: Areas where illicit activity is widespread. (0=No, 1=Yes) [v2elsnlfc_11]
- 12: Areas that are very sparsely populated. (0=No, 1=Yes) [v2elsnlfc_12]
- 13: Areas that are remote (difficult to reach by available transportation, for example). (0=No, 1=Yes) [v2elsnlfc_13]
- 14: Areas where there are indigenous populations. (0=No, 1=Yes) [v2elsnlfc_14]
- 15: Areas where the national ruling party or group is strong. (0=No, 1=Yes) [v2elsnlfc_15]
- 16: Areas where the national ruling party or group is weak. (0=No, 1=Yes) [v2elsnlfc_16]
- 17: Areas that were subject to a longer period of foreign rule. (0=No, 1=Yes) [v2elsnlfc_17]
- 18: Areas that were subject to a shorter period of foreign rule. (0=No, 1=Yes) [v2elsnlfc_18]
- 19: Areas that were recently subject to foreign rule. (0=No, 1=Yes) [v2elsnlfc_19]
- 20: Areas that have not recently been subject to foreign rule. (0=No, 1=Yes) [v2elsnlfc_20]
- 21: None of the above. (0=No, 1=Yes) [v2elsnlfc_21]

Scale: Series of dichotomous scales.

Answer-type: Multiple selection.

Data release: 1-9.

Cross-coder aggregation: Mean.

Citation: *V-Dem Codebook* (see suggested citation at the top of this document).

Years: 1900-2018

3.1.31 Subnational election area more free and fair name (C) (v2elsnmore)

Project Manager(s): Kelly McMann

Question: In which areas of the country are subnational elections significantly *more* free and fair than the country average for subnational elections?

- The variable tags of the following questions within Direct Democracy have been renamed:
 - * Obligatory referendum super majority (v2ddspmjor) to (v2ddspmor)
 - * Plebiscite permitted (v2ddlegpl) to (v2ddlexpl)
 - * Plebiscite participation threshold (v2ddbndpl) to (v2ddpartpl)
 - * Plebiscite approval threshold (v2ddthrcn) to (v2ddapprpl)
 - * Plebiscite administrative threshold (v2dddistspl) to (v2ddadmpl)
 - * Plebiscite super majority (v2ddspmpl) to (v2ddspmpl)
 - * Initiatives permitted (v2ddlegci) to (v2ddlexci)
 - * Initiatives signatures (v2ddsigcin) to (v2ddsignci)
 - * Initiatives signatures % (v2ddsigcip) to (v2ddsigpci)
 - * Initiatives signature-gathering time limit (v2ddgrtlci) to (v2ddsiglci)
 - * Initiatives signature-gathering period (v2ddgrgpci) to (v2ddsigdci)
 - * Initiatives participation threshold (v2ddbndci) to (v2ddpartci)
 - * Initiatives approval threshold (v2ddthreci) to (v2ddapprci)
 - * Initiatives super majority (v2ddspmjci) to (v2ddspmcici)
 - * Referendums permitted (v2ddlegrf) to (v2ddlexrf)
 - * Referendums signatures (v2ddsigrfn) to (v2ddsignrf)
 - * Referendums signature-gathering limit (v2ddgrtrlrf) to (v2ddsiglrf)
 - * Referendums signature-gathering period (v2ddgrgrprf) to (v2ddsigdrf)
 - * Referendums participation threshold (v2ddbndrf) to (v2ddpartrf)
 - * Referendums approval threshold (v2ddthrerf) to (v2ddapprrf)
 - * Referendums super majority (v2ddspmrjrf) to (v2ddspmrjrf)
 - * Number of popular votes this year (v2ddnumvot) to (v2ddyral)
 - * Occurrence of any type of popular vote this year credible (v2ddvotcrd) to (v2ddcredal)
- The information in the codebook about Male suffrage (v2msuffrage), Female suffrage (v2fsuffrage) and Suffrage (v2asuffrage) has been corrected to indicate that responses are given in percent.
- Income inequality, Gini (e_Unequal_UTIP) has been corrected to measure whether the Gini coefficient is above the mean.

- **Other new or modified information**

- Updated list of countries.
- Updated list of variables.
- Updated citations for V-Dem Reference materials.
- New citation for Ordinal versions of V-Dem’s indices (Lindberg 2016).
- Gaps in coding periods should be interpreted as including the start and end year of the gap. For example, Germany is coded 1900–1945, 1949–2016, which means that the years 1946, -47, and -48 are excluded from the dataset.
- Freedom of expression index (v2x_freexp) was included in Electoral democracy index (v2x_polyarchy) in v6. This has been fixed for v7, where the Expanded freedom of expression index (v2x_freexp_thick) instead is included in the Electoral democracy index.

15.4 New in Version 6 Compared to Version 5

- **Changes:**

- All CSV files are now encoded in UTF-8. When importing CSV data on platforms where UTF-8 is not the default, ensure that the correct encoding is specified in order to properly render diacritics for text variables.
- v2x_suffrage and v2elsuffrage have been modified with small changes to their respective comprising indices.

- **Bug fixes:**

- Superfluous variables corresponding to the means of the binary versions of the following variables have been removed: v2exdfdsgh, v2exdfdsghs, v2exdjcbhg, v2exctlhg, v2exctlhs, v2exrmhsol, v2exrmhgnp.
- Country specific coding periods were previously incorrectly set for the 6.1 country-date dataset. As a result, 279 observations falling outside of the V-Dem coding periods have been removed.
- Missing data for the variables v2expathhg and v2exaphogp for Serbia from 1900 to 1941 have been filled in.
- A bug was fixed regarding the ordinalization of indices. Previously, for the corresponding 5 category ordinal variables, observations that were meant to be 0.75 were incorrectly set as 0.5.

- **New indices and indicators**

- Civil liberties index
- Private civil liberties index
- Physical integrity rights index
- Political civil liberties index
- Additive polyarchy index
- Multiplicative polyarchy index
- Divided party control of legislature index
- Division of power index
- HOG term length by law
- HOS term length by law
- Election voter turnout
- Primary school enrollment
- Secondary school enrollment
- Secondary tertiary enrollment
- Political equality comments

- **Modified indices and indicators**

- Electoral democracy index has been updated with a new aggregation formula.

- **Other new or modified information**

- Providing a standard calculation for standard deviation which is marked with the suffix "sd" (*e.g.*, v2elmulpar_sd). The SD might be used to compute the standard frequentist confidence intervals.
- Providing uniqueness scores to the structure of aggregations – all indices and indicators. Uniqueness is the variance that is ‘unique’ to the variable and not shared with other variables. It is equal to 1 – communality (variance that is shared with other variables). Factor loadings are the weights and correlations between each variable and the factor. The higher the load the more relevant in defining the factor’s dimensionality. A negative value indicates an inverse impact on the factor.
- Updated lists of number of variables.
- Updated list of countries.